International Conference on Innovative Trends in Electronics Communication and Applications 2015 [ICIECA 2015]

ISBN	978-81-929742-6-2		**VOL**	01
Website	icieca.in		**eMail**	icieca@asdf.res.in
Received	02 - April - 2015		**Accepted**	15 - November - 2015
Article ID	ICIECA013		**eAID**	ICIECA.2015.013

An Efficient Motion Based Video Object Detection and Tracking System

G Sharmila Sujatha[1], V Valli Kumari[1]
[1]Research Scholar, Dept., of CS & SE, College of Engineering (A), Andhra University

Abstract: *In Computer Vision, one of the critical tasks is to detect and track objects. Even today, the issues related to this remain an active area of research. Major limitation of the existing monitoring system for performing the automatic detection of moving objects, is the inability in understanding the narrow variation in brightness between the background and the moving object. Therefore, systems fail to perform the detection of moving target. Subsequently, the results obtained will be full of noise and also the computation time will be increased. Recent advances in multimedia and its associated technologies seek for the processing of video clips database. This research proposes object detection and tracking system in which video segmentation, feature extraction, feature clustering and object detection are combined seamlessly using a single feature. The system is comprised of feature extraction phase and will be followed by tracking of similar video clips for the given query clip. The features such as color, texture, edge density and motion are extracted from each frame. In the feature extraction, initially, the motion feature is extracted using an efficient motion estimation algorithm based on similarity measures followed by other feature extractions.*

Keywords: Tracking System, Video Object, Detection**.**

INTRODUCTION

Detection and tracking of moving objects is important in the analysis of video data [5]. Motion is detected using statistical hypothesis testing on difference image between two consecutive frames [2]. Tracking-based approaches for abandoned object detection often become unreliable in complex surveillance videos due to occlusions, lighting changes, and other factors [13]. Video surveillance has been widely used in recent years to enhance public safety and privacy protection. A video surveillance system that deals with content analysis and activity monitoring needs efficient transmission and storage of the surveillance video data. Video compression techniques can be used to achieve this goal by reducing the size of the video with no or small quality loss [12]. A distributed video-surveillance system for the detection of dangerous situations related to the presence of abandoned objects in the waiting rooms of unattended railway stations is presented [1]. Autonomous video surveillance and monitoring has a rich history [10].

In general video surveillance systems uses background estimation and subtraction for the detection and tracking of moving objects. Real world applications can deliver a better performance provided if they are able to tolerate the presence of outliers in the data [7]. In the recent years both in home and business environments for security and management of access points, automatic visual object counting and video surveillance have gained important application [14]. Motion analysis algorithms are based on

Cite this article as: G Sharmila Sujatha, V Valli Kumari. "An Efficient Motion Based Video Object Detection and Tracking System". *International Conference on Innovative Trends in Electronics Communication and Applications (2015)*: 85-97. Print.

processing of multiple-regression pseudo spectrums [15]. Identifying the significant features for performing accurate classification is known as feature selection. The major phases in finding out the feature subset selection system are using best search strategy for finding features. Subsequently, evaluation for testing the significance of these features identified in previous phase [4].In many applications for performing traffic monitoring, human motion capture and also for video surveillance using the moving object detection using background subtraction methods [3].

Detected objects are tracked, and each tracked object has a state for handling occlusion and interference [6]. In the recent past, autonomous video surveillance and monitoring has gained lot of importance as they are able to track human motion in indoor and controlled outdoor environments efficiently [9]. Detection of moving objects is done based on multi-layer lidar that also characterize a zone of interest for reducing the computation complexity of the perception process. Then, fixed frame is used for objects localization and tracking [11]. The Sequence Frame Detection Accuracy (SFDA) is a frame-level measure that penalizes for fragmentations in the spatial dimension while accounting for number of objects detected, missed detects, false alarms and spatial alignment of system output and ground truth objects [8]. Tracking-based approaches for abandoned object detection often become unreliable in complex surveillance videos due to occlusions, lighting changes, and other factors [10].

The rest of the paper is organized as follows. Section II provides various researches conducted in relation to our proposed work. Section III explains about the design strategy and the proposed method. Section IV shows the result and discussion of our proposed method and finally section V concludes our proposed method for video based motion detection and tracking.

Review of recent researches

A numerous researches have been presented in the literature for the detection and tracking of moving object in videos. Almost all existing methods for static suspicious object detection are expected at finding abandoned objects using a static camera in a public place. A brief review of some recent researches is presented here.

Moving objects detection and tracking in video stream were basic fundamental and critical tasks in many computer vision applications. R. Bogush *et al.* [16] proposed an effectiveness increase of algorithms for moving objects detection and tracking. For that, they used additive minimax similarity function. Background reconstruction algorithm was developed. Moving and tracking objects detection algorithms were modified on the basis of additive minimax similarity function.

Vehicle tracking and detection plays an important role in traffic surveillance, still a crucial task in many applications. Template matching was one of the methods used for vehicle detection and tracking. There were several researchers and developers worked on that area. Robust and reliable vehicle detection was a critical step of vehicle recognition. Rajiv Kumar Nath *et al* [17] presented a review of recent template matching methods for detection and tracking of vehicle. Their focus was on systems where the camera was mounted on the vehicle and being fixed such as in traffic/driveway monitoring systems. They discussed the general template matching followed by problem of on-road vehicle detection using template matching.

Shih-Chia Huang [18] has proposed a novel and accurate approach for motion detection in the automatic video surveillance system. In which the methods comprises of modules namely background modelling (BM) module, an alarm trigger (AT) module, and an object extraction (OE) ,which effectively detect the moving objects. In BM module, a novel two phase background approach was adopted which uses a raped matching fallowed by accurate matching for generating optimum background pixels. AT module removes the irrelevant examination of entire background region, thereby facilitates the subsequent OE module to process only the blocks containing moving objects. In last, the OE module was formed as the binary object detection mask for accurate detection of moving objects. The results obtained by the proposed (PRO) method were qualitatively and quantitatively analyzed and compared to the results obtained by other state-of-the-art methods. The observation show that this proposed PRO method outperformed other methods by an F1 metric accuracy rate of up to 53.43%.

Jalil Rasekhi *et al.* [19] has given an detailed explanation of automatic system for airplane detection and tracking using Support Vector Machine (SVM) and the wavelet transform . An experiment was performed to recognize airplane in the opening frame of a video sequence in which 50 airplane images in different situations are used. SVM classifier uses a vector of features for classification of objects pixels and background pixels. Thus the learned model has the ability to detect the airplane in original videos as well as novel images. Therefore, the system acts as a tracker for original videos and as a interpreter for novel images.

After airplane detection in the first frame, the feature vectors of that frame were used to train the SVM classifier. For new video frame, SVM was applied to test the pixels and form a confidence map. Daubechies 4th level of wavelet coefficients corresponding

Cite this article as: G Sharmila Sujatha, V Valli Kumari. "An Efficient Motion Based Video Object Detection and Tracking System". *International Conference on Innovative Trends in Electronics Communication and Applications (2015)*: 85-97. Print.

to input image were used as features. Conducting simulations, it was demonstrated that airplane detection and tracking based on wavelet transform and SVM classification result in acceptable and efficient performance.

YingLi Tian *et al.* [20] have presented a framework for robustly and efficiently detecting abandoned and removed objects based on background subtraction (BGS) and fore-ground analysis with complement of tracking to reduce false positives. In the system background was modelled by three Gaussian mixtures. A person-detection process was also integrated to distinguish static objects from stationary people. They have tested the robustness and efficiency of the proposed method on IBM Smart Surveillance Solutions for public safety applications in big cities and evaluated by several public databases. The test and evaluation had demonstrated that their method was efficient to run on real-time, while being robust to quick-lighting changes and occlusions in complex environments.

Carlos R. del-Blanco *et al.* [21]have proposed an efficient visual detection and tracking framework for the tasks of object counting and surveillance, which meets the requirements of the consumer electronics: off-the-shelf equipment, easy installation and configuration, and unsupervised working conditions. This work was accomplished by a novel Bayesian tracking model that can manage multimodal distributions without explicitly computing the association between tracked objects and detections. It was robust to erroneous, distorted and missing detections.

With the advancement of MEMS technologies, sensor networks have opened up broad application prospects. An important issue in wireless sensor networks was object detection and tracking, which typically involves two basic components, collaborative data processing and object location reporting. The former aims to have sensors collaborating in determining a concise digest of object location information, while the latter aims to transport a concise digest to sink in a timely manner. That issue has been intensively studied in individual objects, such as intruders. However, the characteristic of continuous objects has posed new challenges to this issue. Continuous objects can diffuse, increase in size, or split into multiple continuous objects, such as a noxious gas. Shin-Chih Tu *et al.* [22] proposed a scalable, topology-control-based approach for continuous object detection and tracking. Extensive simulations were conducted, which showed a significant improvement over existing solutions.

Hui Kong *et al*. [23] presented a novel framework for detecting non-flat abandoned objects by matching a reference and a target video sequence. The reference video was taken by a moving camera when there was no suspicious object in the scene. The target video was taken by a camera following the same route and may contain extra objects. The objective was to find these objects. GPS information was used to roughly align the two videos and find the corresponding frame pairs. Based on the GPS alignment, four simple but effective ideas were proposed to achieve the objective: an inter-sequence geometric alignment based on homographies, which was computed by a modified RANSAC, to find all possible suspicious areas, an intra-sequence geometric alignment to remove false alarms caused by high objects, a local appearance comparison between two aligned intra-sequence frames to remove false alarms in flat areas, and a temporal filtering step to confirm the existence of suspicious objects.

Peter Dunne *et al.* [24] presented an approach to localized object detection that was not dependent upon background image construction or object modeling. It was designed to work through camera embedded software using spare processing capacity in a visual signal processor. It used a localized temporal difference change detector and particle filter type likelihood to detect possible trackable objects, and to find a point within a detected object at which a particle filter tracker might be initialized.

Subhabrata Bhattacharya et al. [25] an overview of these problems and the associated limitations of some of the conventional techniques typically employed for these applications. They began with a study of various image registration techniques that were required to eliminate motion induced by the motion of the aerial sensor. Next, they presented a technique for detecting moving objects from the ego-motion compensated input sequence. Finally, they described a methodology for tracking already detected objects using their motion history.

Kalpesh R Jadav, Prof.M.A.Lokhandwal and Prof.A.P.Gharge[29] ,"Vision based moving object detection and tracking", used in video surveillance system. This monitor system cannot find the difference of brightness between the backgrounds and moving object is so small. The accurate moving objet is cannot be achieved. The computation time also increase to heal the noise. The target identified in video clip is on rough area.

PROPOSED METHODOLOGY FOR DESIGNING AN EFFICIENT MOTION BASED VIDEO

Motion Estimation for object detection and tracking

Motion estimation for object detection and tracking has been a major area of research in the recent years. The motion estimation in object tracking and detection plays a major role in estimating the correct object to be tracked. Motion estimation, which generally means the computation of velocity of the moving object in a sequence has been a major problem in image processing and gained more importance in the recent years. However, if the camera is moving, it becomes a difficult task since the image motion is generated by the combined effects of camera motion, structure, and object motion. The intention of moving detection is to extract changing region from image sequence. At present common methods of moving target detection have frame difference, background subtraction, optical flow and motion energy.

Steps involved in our proposed method

The proposed method of motion based video object detection and tracking system consists of the processes such as Video Segmentation, Feature Extraction and Tracking. The first step in our approach is to segment the database video clips into different frames or shots. Next is the Feature extraction step. In our proposed method we extract various features from the segmented image such as color feature, edge density feature and texture feature. Along with these features, motion feature is also extracted which is the major consideration in our proposed approach. The flow diagram for our proposed method shown below,

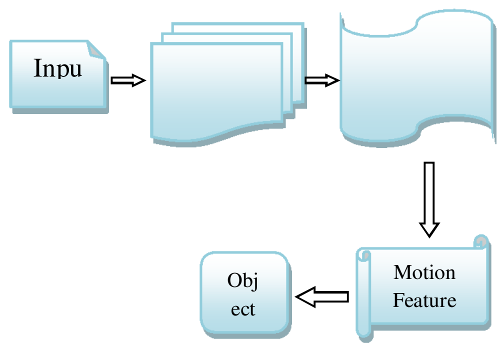


Fig 1: Flow diagram for proposed motion based object detection and tracking

Shot Segmentation

A large amount of digital videos have been generated due to the rapid development of computing and network infrastructures. Normally videos can be represented by a classified structure, while shots are the basis units for constructing high level semantic scenes. The shot boundary recognition is an important preprocessing step for efficient rumination and further content evaluation. A shot consists of succeeding frames which are usually captured by a single camera action. Typically, there are no significant content changes between successive frames in the shot [27]. Once the shot segmentation is done the next step in our process is feature extraction. The shot segmentation is performed with the help of block matching technique which is defined as follows,

Step1. Split the frame into blocks of block size 16.

Step2. First perform block matching for two blocks which are adjacent to eac other.

Step2.1. Then in block matching find the distance of blocks.

Step2.2. Take mean value of distance.

Step3. Repeat these steps for each frame in the video.

Step4. Check whether the mean value is greater than 1.

Step5. The mean value greater than 1 is stored in a variable and these variable and these variable lengths are measured.

Thus the video are shot segmented into various frames with block matching technique.

Cite this article as: G Sharmila Sujatha, V Valli Kumari. “An Efficient Motion Based Video Object Detection and Tracking System”. *International Conference on Innovative Trends in Electronics Communication and Applications (2015)*: 85-97. Print.

Segmentation of Shots using Fuzzy C Means Clustering

Cluster analysis is a technique for classifying data, i.e., to divide a given dataset into a set of classes or clusters. The goal is to divide the dataset in such a way that two cases from the same cluster are as similar as possible and two cases from different clusters are as dissimilar as possible. The idea behind cluster repulsion is to combine an attraction of data to clusters with repulsion between different clusters. Here, the distance between clusters and the data points assigned to them should be minimized. Fuzzy c-means (FCM)is a technique of clustering which permits one piece of data to two or more clusters.

The degrees of membership to which a given data point belongs to the different clusters are computed from the distances of the data point to the cluster centers with respect to the size and the shape of the cluster as stated by the additional prototype information. The closer a data point lies to the center of a cluster, the higher is its degree of membership to this cluster. Hence the problem to divide a dataset $D=\{\vec{d}_1,\vec{d}_2,...,\vec{d}_n\}\subseteq R^P$ into k clusters can be stated as the task to minimize the distances of the data points to the cluster centers and to maximize the degrees of membership.

In probabilistic fuzzy clustering the task is to minimize the objective function:

$$f(D,\delta,\psi)=\sum_{i=1}^{k}\sum_{j=1}^{n}\mu_{ij}^{z}t^{2}(\vec{\lambda}_i,\vec{d}_j)$$

Where $\mu_{ij}\in[0,1]$ is the membership degree of datum $\vec{d}_j$ to cluster k_i, λ_i is the prototype of cluster k_i , and $t\ (\vec{\lambda}_i,\vec{d}_j)$ is the distance between datum $\vec{d}_j$ and prototype λ_i. ψ is the set of all k clusters $\vec{\lambda}_1,\vec{\lambda}_2,...,\vec{\lambda}_k$. $\delta=[\mu_{ij}]$ is called the fuzzy partition matrix and the parameter z is called the fuzzifier.

In possibilistic fuzzy clustering, more intuitive assignment of degrees of membership is achieved by dropping constraint which forces μ_{ij} away from zero for all $i\in\{1,2,...,k\}$. That is, the objective function J is modified to

$$f(D,\delta,\psi)=\sum_{i=1}^{k}\sum_{j=1}^{n}\mu_{ij}^{z}t^{2}(\vec{\lambda}_i,\vec{d}_j)+\sum_{i=1}^{k}\sigma_i\sum_{j=1}^{n}(1-\mu_{ij})^{z}$$

Where, $\sigma_i>0$.

The first term leads to a minimization of the weighted distances while the second term suppresses the trivial solution by making $\mu_{ij}\neq 0\ for\,all\ i\in\{1,2,...,k\}$. This approach is called possibilistic clustering, because the membership degrees for one datum resemble the possibility. The formula for updating the membership degrees that is derived from this objective function is:

$$\mu_{ij}=\frac{1}{1+\left(\frac{t^{2}(\vec{d}_j,\vec{\lambda}_i)}{\sigma_i}\right)^{\frac{1}{z-1}}}$$

σ is chosen for each cluster separately and can be determined by the equation:

$$\sigma_j=\frac{B}{R_i}\sum_{j=1}^{n}\mu_{ij}{}^{z}t^{2}(\vec{d}_j,\vec{\lambda}_i)$$

where usually, $B=1$ and $R_i=\sum\mu_{ij}$.

After the segmentation process the feature extraction process is performed where various features from the segmented frames are extracted for tracking process.

Feature extraction

When the input data to an algorithm is too large to be processed and it is suspected to be notoriously redundant then the input data will be transformed into a reduced representation set of features. The conversion of input data into a set of features is called feature extraction [26]. Extraction of image features and use of these features to represent image visual content is normally termed as feature extraction. Feature extraction involves reducing the amount of resources required to describe a large set of data accurately.

Cite this article as: G Sharmila Sujatha, V Valli Kumari. "An Efficient Motion Based Video Object Detection and Tracking System". *International Conference on Innovative Trends in Electronics Communication and Applications (2015)*: 85-97. Print.

Color Feature Extraction

The HSV (Hue, Saturation, Value) color components are more related to human perception and hence the color histogram extraction is based on HSV color space. Normally based on S component, the color quantization in HSV color space separate gray bins from others and divides the other equally.Fig 1(a) shows the partition in red SV plane [26]. Inspite of this, the nature of HSV color space is that the colors of low V value looks more alike than colors of high V value with respect to different saturations. In the below example the color similarity between A7 to A9 (normally black) is greater than A4 to A6

The cylindrical HSV space is converted into cone space to solve the above problem.The cylindrical HSV point $A(H_a, S_a, V_a)$ is related to conical HSV point $B(H_b, S_b, V_b)$. The transformation is given by

$$\begin{cases} \mathrm{H}_b = \mathrm{H}_a \\ \mathrm{S}_b = S_a \cdot V_a \\ V_b = V_a \end{cases} \qquad \ldots\ldots\ldots \quad (1)$$

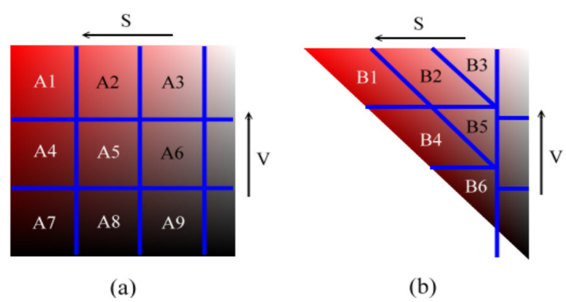


Fig2:Color quantization in SV plane (a) Cylindrical space quantization (b) Conical Space quantization

Comparing fig 2(a) and fig2(b), the color of A7 to A9 correspond to B6 and gray bins.This improves the color quantization of dark colors in HSV color space along with the reduction of number of bins used.

Edge density feature extraction

Edge density may be defined as the quality of an image that can point out the regions by means of the magnitude of the edge of the object available in the image. First, the image region is resampled therefore extracting the feature. The resampled regions of the image are administered to gray scaling operation so that each region of the segmented image that is in RGB color space is transformed to grayscale. In our proposed method we have utilized canny edge detection for feature extraction. The Canny algorithm uses an optimal edge detector based on a set of criteria which include finding the most edges by minimizing the error rate, marking edges as closely as possible to the actual edges to maximize localization, and marking edges only once when a single edge exists for minimal response. The canny algorithm consists of various steps like smoothing, finding gradients, non-maximum suppression, double thresholding and edge tracking by hysteresis.

Texture Feature Extraction

The texture features from the image region are extracted by building a colour texture histogram by using a Local Binary Pattern (LBP). To summarize local gray level structure of the image LBP operator can be employed. It is defined as a gray scale invariant texture measure, derived from a common definition of texture in a local neighbourhood

Local Binary Pattern

The local spatial structure of the image texture are characterized by a gray-scale texture operator called Local Binary Pattern (LBP).LBP indicates a magnitude relation between a center pixel and its neighbouring pixels in a micropattern. Given a central pixel in the image, a pattern code is computed by comparing its value with those of its neighborhoods using the expression given below,

Cite this article as: G Sharmila Sujatha, V Valli Kumari. "An Efficient Motion Based Video Object Detection and Tracking System". *International Conference on Innovative Trends in Electronics Communication and Applications (2015)*: 85-97. Print.

$$LBP_{N,R}(x_i, y_j) = \sum_{N=0}^{N-1} F(P_R - P_C)2^N. \quad (4)$$

where P_C is the gray value of the central pixel, P_R is the value of its neighbour, N is the total number of neighbours involved and R is the radius of the neighborhood. The eqn (4) means that the differences in the neighborhood are derived as N bit binary number that results in 2^N distinct value for the LBP code. The texture can thus be described as 2^N discrete distribution given by,

$$T \approx t\,(LBP_{N,R}(x_i, y_j)) \quad (5)$$

For calculating the LBP feature vector for a given image of size M x N, the central part alone is considered since large neighborhood cannot be used on the borders. The LBP code is calculated for each pixel. The LBP code may be represented as kernel structure index which can be represented as below,

85	75	126
99	95	141
91	91	100

⇨

0	0	1
1		1
0	0	1

Binary: 00111001 Decimal: 57.

Fig 3: An example kernel index structure for LBP.

As per the above kernel the pixels values are being identified and the corresponding decimal values are found out. The LBP feature extraction methods generate the feature descriptor information for each of these images.

Motion feature extraction for motion estimation in video

Generally motion estimation is the process of detecting the motion vectors which forms the transition from one frame to other in a video sequence. The object detection and tracking in any video can be detected based on the motion of the object at consecutive sequences and this can help us in detecting the object. In our proposed method we utilize block matching algorithm for motion estimation process.

Block matching algorithm

The major consideration behind the motion estimation is that the patterns which correspond to objects and background in a frame of video sequence move within the frame to form corresponding objects on the subsequent frame. In block matching proc3ess we divide the current frame into a matrix of 'macro blocks. These macro blocks are then compared with equivalent block and its adjacent neighbors in the previous frame to create a vector that specify the movement of a macro block from one location to another in the previous frame. Thus the motion estimated in the current frame can be obtained by calculating the movements of all macro blocks in a frame. The search area in the macro block is normally reserved upto P pixels. This is called the search parameter and this will be larger for larger motion which in turn takes more execution time. Normally the macro block is a square of side 16 pixels and the search parameter P is 7 pixels. The general block matching idea is shown in fig 4.

The matching of this macro block is based on the cost function of the various macro blocks. The macro blocks whose output cost function is the least cost is the one that matches closest to the current block. There is various cost function available and the least expensive cost function like Mean Square Error (MSE) is selected for the cost calculation in our method. Based on these MSE values we match the blocks. The expression for the Mean Square Error (MSE) is given the eqn below,

Cite this article as: G Sharmila Sujatha, V Valli Kumari. "An Efficient Motion Based Video Object Detection and Tracking System". *International Conference on Innovative Trends in Electronics Communication and Applications (2015)*: 85-97. Print.

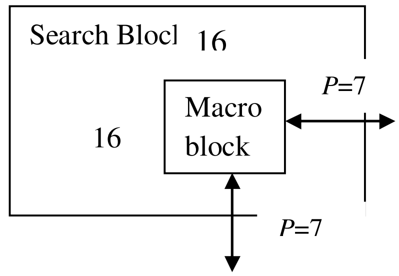


Fig 4: Macro block with side 16 and search parameter P=7

$$\text{Mean Square Error (MSE)} = \frac{1}{N^2}\sum_{p=0}^{N-1}\sum_{q=0}^{N-1}(P_C - P_R)^2. \qquad (7)$$

Here N is the number of sides in macro block, P_C and P_R are the pixels that are being compared from the current block and the reference block

Moving Object Detection and Tracking

From each frame the moving objects are detected. There are different basic methods employed for moving object detection like

- Background Subtraction.
- Temporal Differencing.
- Optical Flow.

The background subtraction method is a straightforward method for moving target detection. In background subtraction method, it is assumed that the background is static so that the background does not change with the number of frames. First the difference between the object P and the background Q is calculated using the formula,

$$D(a,b) = \left|P(a,b) - Q(a,b)\right| \qquad (8)$$

Now threshold the difference using the formula given below,

$$Z(a,b) = \begin{cases} 1, & D(a,b) > T \\ 0, & otherwise \end{cases} \qquad (9)$$

The threshold can be chosen using the gray histogram by taking the bottom value between the two peaks as the threshold[28]. Basically Object tracking is used to find the location of the target in different frames in a sequence of images. The main work of target tracking is to select good target characteristics and use appropriate searching methods. We employ the Image Difference Algorithm for Moving Object Detection and Tracking in our work.

Object Tracking

To achieve a good detection rate on each shot of a video frame, the detection and tracking were combined and some rules are formed to get a complete tracking process. The tracking is of two types,

1) Forward Tracking.
2) Backward Tracking.

Cite this article as: G Sharmila Sujatha, V Valli Kumari. "An Efficient Motion Based Video Object Detection and Tracking System". *International Conference on Innovative Trends in Electronics Communication and Applications (2015)*: 85-97. Print.

Forward Tracking

The forward tracking process is performed on each frame, beginning from frames where the object have been detected. While tracking, same object may be detected many times in a shot which can result in multiple tracking of the object, which may result in over time consumption. Inorder to overcome this problem, some tracking rules are used to identify whether the detected objects are multiplied or not. The rule is generally based on the percentage of overlap between the detected object and the one from the forward tracking in the same frame which is represented as follows, $$\mathrm{F_T} = \max_j \frac{S_{(F_T \cap W_j)}}{\min\left(S_{W_j}, S_{F_T}\right)}$$

(10)

where S_{W_j} is the area of the j^{th} detection and S_{F_T} is the area of the forward tracking. Also $S_{(F_T \cap W_j)}$ represents the area recovered by the detection process.

Backward Tracking

Backward Tracking is performed on each frame to provide an additional set of object being tracked. The backward tracking is very useful in case the object is not detected at the beginning but in the center of the frame. The forward tracking usually represents the object tracking from the detected frame to the end the complete shot while backward tracking gives the unnoticed result from the first frame of the shot to the frame in which the last object detection has performed. Further the backward tracking can also proves to be effective when the forward tracking fails to locate the position of the object in a particular frame. This may be due to occlusion, bad illumination or due to tracker sticks to the background. That is when an object in a frame1 is not tracked correctly and the same object is tracked in frame 5, the information will be propagated back and will provide tracking of the object in the first frame.

Object Tracking Process

In the proposed method, the tracking process is done by comparing the features extracted in the present frame with the extracted features in the previous frame. Suppose we have a total of M number of extracted feature in present frame and N number of feature in the previous frame. Hence a total of M x N matching is required. In this experiment we utilized Euclidean Distance for this matching process. Suppose we consider two feature vectors v_i and vj where $i = 1,2,.....M$ and $j = 1,2,......N$,then the Euclidean distance is calculated by using the below equation,

$$D\left(\mathrm{v}_i, v_j\right) = \sqrt{\frac{1}{d}\sum_{m=1}^{i}\left(v_{im} - v_{jm}\right)^2} \quad (11)$$

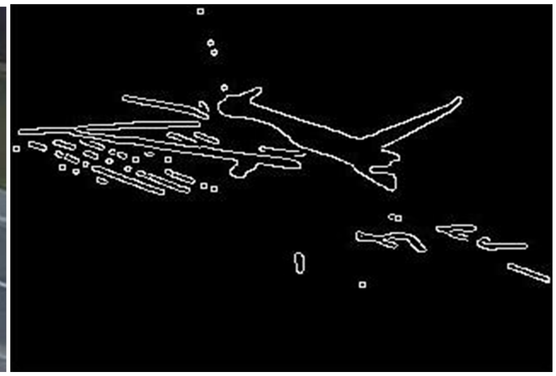
Where d is the dimension of the feature vector selected. Once the distances between the features are calculated, the minimum distance object is tracked. This object is the object which we have to be tracked from the video clip. This process is then carried out for different shots and the movements of the object is identified which helps in efficient tracking of the object.

RESULTS AND DISCUSSION

The proposed motion based object detection and tracking system was implemented in the working platform of MATLAB. The detection and tracking process is tested with different frames of video and the upcoming result of the proposed work has been shown below. Initially, the video are segmented to different shots or frames and then features are extracted followed by the detection and tracking process. The results obtained by our proposed method for different frames of the input video is shown below,

This feature extraction provides better outcome when comparetracked image of the object in the first frame. Similarly for different frames the process is repeated and finally the object is tracked. The proposed methodology proved to be more effective and accurate in object detection and tracking.

Cite this article as: G Sharmila Sujatha, V Valli Kumari. "An Efficient Motion Based Video Object Detection and Tracking System". *International Conference on Innovative Trends in Electronics Communication and Applications (2015)*: 85-97. Print.

(a) Input frame (b) feature extracted output


(c)Detected object d) Object tracking

Fig 5:Results of object tracking in first frame (Video).

PERFORMANCE ANALYSIS

The values obtained from the calculation are given in Table 1. These values are used for the analysis of performance between the proposed and existing method.

Table1: Precision and Recall for the proposed method

S.No	Performance Analysis				F-Measure	
	Precision		Recall			
	Proposed Method	Existing Method	Proposed Method	Existing Method	Proposed Method	Existing Method
1	0.9988	0.71	1	0.15	0.9993	0.2477
2	0.9995	0.65	1	0.24	0.9997	0.3506
3	0.9995	0.59	1	0.35	0.9997	0.4394
4	0.9935	0.52	1	0.46	0.9967	0.4882
5	0.9925	0.47	1	0.52	0.9962	0.4937

Cite this article as: G Sharmila Sujatha, V Valli Kumari. "An Efficient Motion Based Video Object Detection and Tracking System". *International Conference on Innovative Trends in Electronics Communication and Applications (2015)*: 85-97. Print.

Here the existing method is the previous paper where object detection and tracking using low level features is performed. Each values relating to the methods are entered in the table for comparison and from the table it is clear that our proposed method delivers better precision and recall than the existing method. Here the existing method is the vision based object detection and tracking [29].

The average F-measure value for the proposed and existing method is found out and the corresponding graph is shown in fig 8.

Table 2: Average F-measure for proposed and existing methods

F-measure	Methods	
	Proposed	Existing
Average F-measure	0.9983	0.4039

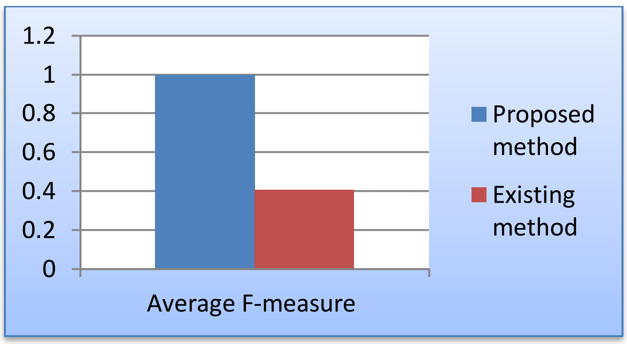


Fig 6: Graphical representation of Average F-measure for proposed and existing method.

CONCLUSION

In this paper we have proposed efficient motion based object detection and tracking system. We developed a unique method using Block matching algorithm and our method proves it is a well-organized method for motion estimation to track and detect the objects in the video. As the results shows the proposed methodology proved to be more efficient and accurate in object detection and tracking than the previous methods. To prove the effectiveness of our proposed method we have compared the precision and recall value along with F-measure of the proposed method with existing method for the object detection and tracking process. As per the performance analysis, it is clear that our proposed method provides better F-measure value when comparing with other method. As a result it can be concluded that our proposed method is efficient in the field of object detection and tracking.

REFERENCES

[1] Claudio Sacchi and Carlo S. Regazzoni, “A Distributed Surveillance System for Detection of Abandoned Objects in Unmanned Railway Environments”, IEEE Transactions on Vehicular Technology, Vol. 49, No. 5, September 2000.

[2] Sang Won Hwang, Eun Yi Kim, Se Hyun Park and Hang Joon Kim, “Object Extraction and Tracking Using Genetic Algorithm”, IEEE Transactions, 2001.

[3] Rita Cucchiara, Costantino Grana, Massimo Piccardi and Andrea Prati, “Detecting Moving Objects, Ghosts and Shadows in Video Streams”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, pp. 1337-1342, October 2003.

[4] Zehang Suna , George Bebisa and Ronald Millerb, “Object detection using feature subset selection”, Journal of the Pattern Recognition Society, March 2004.

Cite this article as: G Sharmila Sujatha, V Valli Kumari. “An Efficient Motion Based Video Object Detection and Tracking System”. *International Conference on Innovative Trends in Electronics Communication and Applications (2015)*: 85-97. Print.

[5] Chandrika Kamath, Abel Gezahegne, Shawn Newsam and George M. Roberts," Salient Points for Tracking Moving Objects in Video", Image and Video Communications and Processing, 2005.

[6] Masayuki Yokoyama and Tomaso Poggio, "A Contour-Based Moving Object Detection and Tracking", In proceedings of Visual Surveillance and Performance Evaluation of Tracking and Surveillance, 2005.

[7] Codrut Ianasi, Vasile Gui, Corneliu I. Toma, and Dan Pescaru, "A Fast Algorithm for Background Tracking in Video Surveillance, Using Nonparametric Kernel Density Estimation", Electronics and Energetics, Vol. 18, No. 1, pp. 127-144, 2005.

[8] Vasant Manohar, Padmanabhan Soundararajan, HarishRaju, Dmitry Goldgof, Rangachar Kasturi, and John Garofolo, "Performance Evaluation of Object Detection and Tracking in Video", In Proceedings of 7th Asian Conference on Computer Vision, Hyderabad, pp. 151–161, 2006.

[9] Kaiqi Huanga, Liangsheng Wanga, Tieniu Tana, and Steve Maybank, "A real-time object detecting and tracking system for outdoor night surveillance", ELSEVIER Pattern Recognition, Vol. 41, pp. 432-444, 2008.

[10] Limin Liu, Zhen Li, Member, and Edward J. Delp, "Efficient and Low Complexity Surveillance Video Compression Using Backward Channel Aware Wyner-Ziv Video Coding", IEEE Transactions, June 2008.

[11] Sergio A. Rodríguez F, Vincent Frémont, Philippe Bonnifait and Véronique Cherfaoui, " An Embedded Multi-Modal System for Object Localization and Tracking", IEEE Intelligent Vehicles Symposium, 2010.

[12] Sherin M. Youssef, Meer A. Hamza and Arige F. Fayed, "Detection and Tracking of Multiple Moving Objects with Occlusion in Smart Video Surveillance Systems", IEEE Transactions, 2010.

[13] Yuping Lin, Qian Yu and Gérard Medioni, "Efficient detection and tracking of moving objects in geo-coordinates", Machine Vision and Applications, Vol. 22, pp. 505–520, 2011.

[14] Thi Thi Zin, Pyke Tin, Takashi Toriu and Hiromitsu Hama, "A Probability-based Model for Detecting Abandoned Objects in Video Surveillance Systems", Proceedings of the World Congress on Engineering, July 2012.

[15] Boris Vishnyakov, Yury Vizilter and Vladimir Knyaz, "Spectrum-Based Object Detection And Tracking Technique For Digital Video Surveillance", International Archives of the Photogrammetry, September 2012.

[16] R. Bogush, S.Maltsev, A. Kastryuk, N. Brovko and D. Hlukhau, "Motion detection and tracking algorithms in video streams", Journal of Science, Mathematics – Physics, Vol. 25, pp. 143-151, 2009

[17] Jalil Rasekhi, MohammadReza Karami and Mojtaba Bandarabadi, "Wavelet Transform and Supervised Learning Methods for Object Tracking", European Journal of Scientific Research ISSN 1450-216X Vol.41 No.4, pp.626-631, 2010

[18] Shih-Chia Huang, "An Advanced Motion Detection Algorithm with Video Quality Analysis for Video Surveillance Systems", IEEE Transactions on Circuits and Systems For Video Technology, Vol. 21, No. 1, January 2011.

[19] Rajiv Kumar Nath and Dr. Swapan Kumar Deb, "On Road Vehicle/Object Detection And Tracking Using Template", Indian Journal of Computer Science and Engineering, Vol 1 No 2, 98-107, Sept 2010.

[20] YingLi Tian, Rogerio Schmidt Feris, Haowei Liu, Arun Hampapur, and Ming-Ting Sun, "Robust Detection of Abandoned and Removed Objects in Complex Surveillance Videos", IEEE Transactions On Systems, Vol. 41, No. 5, September 2011.

[21] Carlos R. del-Blanco, Fernando Jaureguizar, and Narciso García, "An Efficient Multiple Object Detection and Tracking Framework for Automatic Counting and Video Surveillance Applications", IEEE Transactions on Consumer Electronics, Vol. 58, No. 3, August 2012.

[22] Shin-Chih Tu, Guey-Yun Chang, Jang-Ping Sheu, Wei Li and Kun-Ying Hsieh, "Scalable continuous object detection and tracking in sensor networks", Journal of Parallel and Distributed Computing, Vol. 70, pp. 212-224, 2010

[23] Hui Kong, J. Y Audibert and J. Ponce, "Detecting Abandoned Objects With a Moving Camera", Image Processing, Vol. 19, No. 8, pp. 2201-2210, 2010

Cite this article as: G Sharmila Sujatha, V Valli Kumari. "An Efficient Motion Based Video Object Detection and Tracking System". *International Conference on Innovative Trends in Electronics Communication and Applications (2015)*: 85-97. Print.

[24] Peter Dunne and Bogdan J. Matuszewski, “Histogram-based Detection of Moving Objects for Tracker Initialization in Surveillance Video”, International Journal of Grid and Distributed Computing, Vol. 4, No. 3, pp. 71-78, Sept 2011

[25] Subhabrata Bhattacharya, Haroon Idrees, Imran Saleemi, Saad Ali, and Mubarak Shah, “Moving Object Detection and Tracking in Forward Looking Infra-Red Aerial imagery”, Machine Vision Beyond Visible Spectrum, Vol. 1, No. 3, pp. 221-252, 2011

[26] J. Fernandeza, R. Guerreroa, N. Mirandaa and F. Piccolia ,"Multi-Level Paralelism In Image Identification,"Mecanica Computational,Vol.28,pp.227-240,Argentina,Nov 2009.

[27] Li Li, Xianglin Zeng,Xi Li,Weiming Hu and Pengfei Zhu ,"Video Shot Segmentation Using Graph-based Dominant-Set Clustering," In.proc.of. First International Conference on Internet Multimedia Computing and Service,New York,USA,2009.

[28] Liang Xiao and Tong-qiang Li,"Research on Moving Object Detection and Tracking,"In.proc.of.Seventh International Conference on Fuzzy Systems and Knowledge Discovery, pp.2324-2327, 2010.

[29] Kalpesh R Jadav, Prof.M.A.Lokhandwala and Prof.A.P.Gharge,"Vision based moving object detection and tracking",National Conference on Recent Trends in Engineering & Technology,May 2011.

Cite this article as: G Sharmila Sujatha, V Valli Kumari. “An Efficient Motion Based Video Object Detection and Tracking System”. *International Conference on Innovative Trends in Electronics Communication and Applications (2015)*: 85-97. Print.